

Robust Lattice Alignment for K -user MIMO Interference Channels with Imperfect Channel Knowledge

[†]Huang Huang, *Student Member, IEEE*, [†]Vincent K. N. Lau, *Senior Member, IEEE*,

^{*}Yinggang Du and ^{*}Sheng Liu

[†]ECE Department, Hong Kong University of Science and Technology, Hong Kong

^{*}Huawei Technologies, Co. Ltd. China

Abstract—In this paper, we consider a robust lattice alignment design for K -user quasi-static MIMO interference channels with imperfect channel knowledge. With random Gaussian inputs, the conventional interference alignment (IA) method has the feasibility problem when the channel is quasi-static. On the other hand, structured lattices can create structured interference as opposed to the random interference caused by random Gaussian symbols. The structured interference space can be exploited to transmit the desired signals over the gaps. However, the existing alignment methods on the lattice codes for quasi-static channels either require *infinite SNR* or *symmetric interference channel coefficients*. Furthermore, perfect channel state information (CSI) is required for these alignment methods, which is difficult to achieve in practice. In this paper, we propose a robust lattice alignment method for quasi-static MIMO interference channels with imperfect CSI at all SNR regimes, and a two-stage decoding algorithm to decode the desired signal from the *structured interference space*. We derive the achievable data rate based on the proposed robust lattice alignment method, where the design of the precoders, decorrelators, scaling coefficients and *interference quantization coefficients* is jointly formulated as a mixed integer and continuous optimization problem. The effect of imperfect CSI is also accommodated in the optimization formulation, and hence the derived solution is robust to imperfect CSI. We also design a low complex iterative optimization algorithm for our robust lattice alignment method by using the existing iterative IA algorithm that was designed for the conventional IA method. Numerical results verify the advantages of the proposed robust lattice alignment method compared with the TDMA, two-stage ML decoding, generalized Han-Kobayashi (HK), distributive IA and conventional IA methods in the literature.

Index Terms—lattice codes, interference alignment, interference channel, MIMO, imperfect CSI

I. INTRODUCTION

Interference is a fundamental bottleneck in wireless systems. This is partially due to the lack of understanding interference from an information theoretic view. For example, the capacity region of the two-user Gaussian interference channels has been an open problem for over 30 years. It was only shown recently that the Han-Kobayashi (HK) achievable region [1] achieves the capacity region to within one bit [2]. Lately, there have been some breakthroughs on the understanding of K -user interference channels. In [3] and [4], the authors propose an *interference alignment* (IA) method to align interference onto a lower dimensional subspace of each receiver so that

the desired signal can be transmitted on the *interference-free dimensions*. The authors show that IA is optimal in the degrees-of-freedom (DoF) sense (which is a high SNR performance measure) and the total capacity of the K -user interference channels is given by $\frac{K}{2} \log(\text{SNR}) + o(\log(\text{SNR}))$ [3]. There are a number of extensions [5], [6] that have studied the application of IA in K -user quasi-static MIMO interference channels. However, the conventional IA method for K -user interference channels requires infinite dimension in time-varying or frequency-selective channels, and has the feasibility problem. For example, it is shown in [7] that conventional IA in quasi-static MIMO (M transmit and N receive antennas) interference channels is not able to achieve a per user DoF greater than $\frac{M+N}{K+1}$. As a result, there is no satisfactory solution for quasi-static MIMO interference channels for large K due to the feasibility problem.

Besides the feasibility problem, conventional IA solutions have assumed Gaussian input symbols and it is unclear whether it is optimal to employ Gaussian input for interference channels. With the Gaussian input symbols, the interference space is random and all the conventional IA methods try to align all interference to a smaller dimension space and utilize the remaining *interference-free* dimensions to transmit the desired signals. While Gaussian inputs (using random codebook argument) are capacity optimal in multi-access and broadcast channels [8], [9], they are not optimal for interference channels [2], [10]. It is revealed in [10] that structured interference is more preferred in interference networks. Instead of giving up the space that is reserved for the interference and transmitting the desired signal on the remaining *interference-free dimensions*, one could also exploit the structured interference space and transmit desired signals over the *gaps*. This motivates the *lattices* based study of interference alignment. In [11], [12], the authors propose a lattice interference alignment method to align the interfering lattices on a common basis but it only works for symmetric SISO interference channels (where all cross links have the same fading coefficients) or a specialized class of 3-user SISO interference channels (where the products of the fading coefficients are assumed to be rational). In [13], [14], the authors propose interference alignment over real line for K -user quasi-static MIMO interference channels with real channel coefficients and demonstrate that the DoF

$= \frac{MN}{M+N} K$ can be achieved. This approach is further extended to the complex channel in [15] for the compound MIMO broadcast channel. In addition to interference channels, lattice codes have also been widely studied in many communication networks [16] such as point-to-point channels and wireless relay networks. Specifically, in [17], the author shows that lattice codes can achieve the capacity of $\frac{1}{2} \log(1+\text{SNR})$ in the point-to-point additive white Gaussian noise (AWGN) channel. In [18], the authors exploit the lattice codes and propose a compute-and-forward method in wireless relay networks. They show that each relay can decode a linear function of messages from multiple source nodes by using lattice codes. A destination node can decode the desired message, given sufficient linear combinations forwarded from the relay nodes.

All these results indicate that we can potentially benefit from the *structured interference* created by lattice codes. However, in order to have a practical method for exploiting the advantage of the structured interference due to lattice transmissions, there are still some key technical challenges to be addressed.

- **How to align the received lattices on a common basis over irrational quasi-static MIMO interference channels:** Creating an ideal structured interference space [11], [12] requires that all the interfering lattices to be received on a common basis at each of the K receivers. However, this is a difficult requirement and is impossible for irrational fading matrices. While [11], [12] study lattice alignment methods for K user quasi-static interference channels, the methods therein only work for *symmetric interference channels* or a specialized class of 3-user *SISO interference channels*. These approaches cannot be used for general non-symmetric irrational interference channels.
- **How to exploit structured interference at finite SNR:** In [13], [14], the authors propose a lattice alignment algorithm for MIMO interference channels, which is optimal in the DoF sense. However, it is not known whether this approach can be applied at finite SNR, which is an important operating regime in practice. In [19], the authors propose an ergodic interference alignment method which could work at finite SNR regime. However, it requires either infinite time or frequency extension and cannot be applied for constant channels.
- **How to ensure robustness due to imperfect knowledge of CSI:** All of the above works assume perfect knowledge of channel state information (CSI) for facilitating interference alignment. In practice, this is not possible and interference alignment performs very poorly with imperfect CSI [20]. It is quite challenging to incorporate robustness with respect to (w.r.t.) imperfect CSI and exploit the structured interference space at the same time.

In this paper, we explore a robust precoder and decorrelator design for K -user general irrational quasi-static MIMO interference channels with imperfect CSI. We propose a *robust lattice alignment* method (which does not have the feasibility problem) for general irrational constant interference channels, and deduce a *two-stage decoding algorithm* to decode the desired signal from the *structured interference space*. We

derive the achievable data rate based on the proposed method and formulate the precoders, decorrelators, scaling coefficients as well as the interference quantization coefficients¹ design as a mixed integer and continuous optimization problem [21]. By utilizing the alternating optimization technique and by exploiting the structured interference, we derive a low complexity algorithm to determine the precoders, decorrelators as well as the interference quantization coefficients. The effect of imperfect CSI is also accommodated in the optimization formulation and hence, the derived solution is robust to imperfect CSI. To illustrate the benefit of the proposed method, we compare the achievable data rates with those of the conventional IA method [3], the distributed IA method based on alternating optimization [5], the TDMA method, the brute-force two-stage maximum likelihood (ML) decoding method and the generalized HK method [1], [2] with random Gaussian input symbols. We show that the proposed method achieves the same DoF as the conventional IA method if the problem is feasible. On the other hand, when the conventional IA is not feasible, our proposed solution can still offer significant performance gain compared with these baselines.

This paper is organized as follows. In Section II, we outline the system model of the K -user quasi-static MIMO interference channels. In Section III, we review some common interference alignment methods in the literature and provide some preliminary discussions on the lattices. In Section IV, we propose the robust lattice alignment method and formulate it as a mixed integer and continuous optimization problem. In Section V, by using the alternating optimization technique, we derive a low complexity solution. In Section VI, we derive closed-form analysis of the proposed method and the baseline methods under specialized channel realizations. The performance of the proposed method and the existing methods under complex channel realizations is illustrated in Section VII. We conclude with a brief summary of the results in Section VIII.

II. SYSTEM MODEL

A. K -user Quasi-Static MIMO Interference Channels

We consider the K -user quasi-static MIMO Gaussian interference channels as illustrated in Fig. 1. Each transmitter, equipped with M antennas, tries to communicate to its corresponding receiver, which is equipped with N antennas. Specifically, we consider a block of T channel symbols. The channel output at the k -th receiver is described as follows:

$$\mathbf{Y}_k = \sum_i \mathbf{H}_{ki} \bar{\mathbf{X}}_i + \mathbf{Z}_k, \quad (1)$$

where, \mathbf{H}_{ki} is the $N \times M$ MIMO complex fading coefficients from the i -th transmitter to the k -th receiver. $\bar{\mathbf{X}}_i$ is the $M \times T$ complex signal vector transmitted by transmitter i . \mathbf{Z}_k is the $N \times T$ circularly symmetric AWGN vector at receiver k . We assume all noise terms are i.i.d. zero mean complex Gaussian with unit variance.

¹Please refer to Section IV for details about these design parameters.

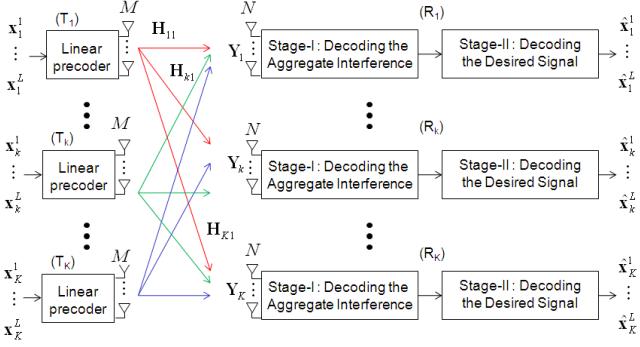


Fig. 1. Quasi-static complex K -user MIMO Gaussian interference channels. Each transmitter, equipped with M antennas, tries to transmit L independent data streams to its corresponding receiver, which is equipped with N antennas.

B. Precoding at Transmitters

In this paper, we assume that each transmitter transmits L independent data streams. Specifically, \mathbf{v}_i^l is the $M \times 1$ precoder for the l -th complex data stream \mathbf{x}_i^l ($1 \times T$) at transmitter i . Therefore, $\overline{\mathbf{X}}_i = \sum_l \mathbf{v}_i^l \mathbf{x}_i^l$. The average power for each data stream is given by $\frac{1}{T} \mathbb{E}[||\mathbf{x}_i^l||^2] = P$. The precoders are chosen such that the total transmit power for each transmitter is no more than γP , i.e.,

$$\sum_{l=1}^L ||\mathbf{v}_i^l||^2 \leq \gamma, \quad \forall i, \quad (2)$$

where $||\mathbf{v}_i^l||^2 = ||\mathbf{v}_i^l||_2^2 = (\mathbf{v}_i^l)^H \mathbf{v}_i^l$.

C. Imperfect Channel State Information Model

The imperfect CSI model can be expressed as:

$$\hat{\mathbf{H}}_{ki} = \mathbf{H}_{ki} + \Delta_{ki}, \quad (3)$$

where $\hat{\mathbf{H}}_{ki}$ is the estimated CSI that is known at the central coordinator. Δ_{ki} is the CSI error satisfying:

$$\begin{aligned} \Delta_{ki} \in \mathcal{E} &= \{ \Delta_{ki} : \|\Delta_{ki}\|_F^2 \leq \epsilon^2 \} \\ &= \{ \Delta_{ki} : \text{Tr}\{\Delta_{ki}(\Delta_{ki})^H\} \leq \epsilon^2 \}. \end{aligned} \quad (4)$$

This deterministic channel estimation error model is widely used in the literature [22]–[24]. Therefore, given the estimated CSI $\hat{\mathbf{H}}_{ki}$, the uncertainty of \mathbf{H}_{ki} is an ellipsoid centered at $\hat{\mathbf{H}}_{ki}$ with radius ϵ .

III. SUMMARY OF THE EXISTING INTERFERENCE ALIGNMENT METHODS AND LATTICES

A. Interference Alignment on a Quasi-Static MIMO Space

It is shown in [3] that in the 3-user quasi-static MIMO interference channels with $M > 1$ antennas, $\frac{3M}{2}$ DoF can be achieved. Specifically, when M is even, each transmitter sends $\frac{M}{2}$ independent streams over the $M \times \frac{M}{2}$ precoder \mathbf{V}_i for transmitter i , i.e.,

$$\overline{\mathbf{X}}_i = \sum_{m=1}^{M/2} \mathbf{v}_i^m \mathbf{x}_i^m = \mathbf{V}_i \mathbf{X}_i. \quad (5)$$

The precoders $\{\mathbf{V}_i\}_{i=1}^3$ are designed in terms of the channel matrices, where the dimension of the interference space is equal to $M/2$ at each receiver. Each receiver can simply cancel the interference by zero-forcing and then decode the desired $M/2$ streams. When M is odd, similar results can be obtained by considering a two time-slot symbol extension of the channel, with the same channel coefficients over the two symbols. However, it is shown in [7] that the above method is not able to achieve a per user DoF greater than $\frac{M+N}{K+1}$.

B. Interference Alignment on a Lattice

In [12], the authors consider IA on a lattice for some special 3-user SISO interference channels ($M = N = 1$). Specifically, the *real* channel fading coefficients have to satisfy the following requirement:

$$\frac{H_{12}}{H_{21}} \times \frac{H_{23}}{H_{32}} \times \frac{H_{31}}{H_{13}} = \frac{p}{q}, \quad (6)$$

where p and q are integers such that $\text{gcd}(p, q) = 1$. $\text{gcd}(\cdot)$ is the greatest common divisor of the integers. Transmitter i chooses the lattice Λ_i generated by using *construction A* described in [25]. In order to align the interfering lattices at each receiver, the lattices are chosen by:

$$H_{12}\Lambda_2 = pH_{13}\Lambda_3, H_{21}\Lambda_1 = qH_{23}\Lambda_3, H_{31}\Lambda_1 = H_{32}\Lambda_2. \quad (7)$$

As a result, a two-stage decoding algorithm can be done at each receiver, which decodes the *aggregate interference* first and then subtracts the aggregate interference from the received signal to decode the desired information. Note that the symmetric channel considered in [11] is a special case in (6). This method requires strong interference channels to decode the interference first (treating the desired signal as additive noise). Furthermore, the requirement of adding the interfering lattice on a common basis at each receiver as in (7) becomes infeasible for general irrational K -user quasi-static interference channels. A simple counter example is given below.

Example 1 (A Simple Counter Example): A simple counter example is given by: $\mathbf{y}_1 = \mathbf{x}_1 + \sqrt{2}\mathbf{x}_2 + \sqrt{3}\mathbf{x}_3 + \mathbf{z}_1$, $\mathbf{y}_2 = \sqrt{5}\mathbf{x}_1 + \mathbf{x}_2 + \sqrt{7}\mathbf{x}_3 + \mathbf{z}_2$, $\mathbf{y}_3 = \sqrt{11}\mathbf{x}_1 + \sqrt{13}\mathbf{x}_2 + \mathbf{x}_3 + \mathbf{z}_3$. It is not possible to align the interfering lattices on a common basis at each of the 3 receivers. ■

C. Interference Alignment over the Real Line

It is shown in [14] that $\frac{MN}{M+N}K$ DoF can be achieved for quasi-static MIMO using IA on the *real* line [13]. This approach is extended to the complex channel in [15] for the compound MIMO broadcast channel. Here we review the basic idea for the SISO interference channels in [13], and it is easy to extend to the MIMO case as in [14], [15]. Specifically, the received signal at receiver k can be represented as [13]

$$y_k = A \left(\sum_{l=0}^{L_k-1} H_{kk} T_{kl} x_k^l + \underbrace{\sum_{i=1, i \neq k}^K \sum_{l=0}^{L_i-1} H_{ki} T_{il} x_i^l}_{I_k} \right) + z_k, \quad (8)$$

where A controls the input power of all users. $x_k^l \in (-Q, Q)_{\mathbb{Z}}$ is one of the integers in the set $(-Q, Q)$, and carries information for the l -th data stream for user k . T_{kl} is a constant real number, which is the direction of the transmitted l -th data stream. It is chosen as monomials with variables from channel coefficients. I_k is the aggregated interference caused by all users. The data streams are aligned if they arrive at the same direction. e.g., interference x_1^l and x_2^l are aligned at receiver 3 if $H_{31}T_1^l = H_{32}T_2^l$. By carefully designing the transmit directions, it is shown in [13], [14] that $\frac{MN}{M+N}K$ DoF can be achieved. However, this method requires infinite SNR and it is not known if this method could be modified for finite SNR regime.

D. Lattices

In this section, we shall review some preliminaries on lattices from [16], [18] and the references therein. A T -dimensional lattice Λ is a set of points in \mathbb{R}^T , and is given in terms of the lattice generator matrix $\mathbf{L} \in \mathbb{R}^{T \times T}$:

$$\Lambda = \{\mathbf{x} = \mathbf{L}\mathbf{w} : \mathbf{w} \in \mathbb{Z}^T\}. \quad (9)$$

A *lattice quantization* is a function, $Q_{\Lambda} : \mathbb{R}^T \rightarrow \Lambda$, that maps a point \mathbf{x} to the nearest lattice point in Euclidean distance:

$$Q_{\Lambda}(\mathbf{x}) = \arg \min_{\lambda \in \Lambda} \|\mathbf{x} - \lambda\|. \quad (10)$$

The *fundamental Voronoi region*, \mathcal{V} , of Λ , is the set of points in \mathbb{R}^T closest to the zero vector, i.e., $\mathcal{V} = \{\mathbf{x} : Q_{\Lambda}(\mathbf{x}) = 0\}$. The modulo- Λ operation w.r.t. the lattice is defined as

$$\mathbf{x} \text{ Mod } \Lambda = \mathbf{x} - Q_{\Lambda}(\mathbf{x}), \quad (11)$$

which is also the quantization error of \mathbf{x} w.r.t. Λ .

A *nested lattice code* \mathcal{L}_1 is the set of all points of a fine lattice Λ_1 that are within the fundamental Voronoi region \mathcal{V} of a *coarse lattice* Λ :

$$\mathcal{L}_1 = \Lambda_1 \cap \mathcal{V} = \{\mathbf{x} : \mathbf{x} = \lambda \text{ Mod } \Lambda, \lambda \in \Lambda_1\}, \quad (12)$$

where Λ is said to be *nested* in Λ_1 , i.e., $\Lambda \subseteq \Lambda_1$. Please refer to Fig. 2 in [18] for an illustration of nested lattice. The rate of a nested lattice code is given by:

$$R = \frac{1}{T} \log(|\mathcal{L}_1|) = \frac{1}{T} \log \frac{\text{Vol}(\mathcal{V})}{\text{Vol}(\mathcal{V}_1)}, \quad (13)$$

where $\text{Vol}(\mathcal{V})$ is the volume of \mathcal{V} .

IV. ROBUST LATTICE ALIGNMENT

In this section, we introduce the framework of our robust lattice alignment method and the associated two-stage decoding algorithm. We derive the minimum achievable data rate and design the precoder, decorrelator and the interference quantization coefficients via an optimization approach.

A. Encoding at Transmitters

In this paper, we adopt the lattice encoding method, which is used both in point-to-point channels [17] and relay networks [18]. Specifically, the data stream \mathbf{x}_k^l is given by:

$$\mathbf{x}_k^l = [\mathbf{t}_k^l - \mathbf{d}_k^l] \text{ Mod } \Lambda + j [\tilde{\mathbf{t}}_k^l - \tilde{\mathbf{d}}_k^l] \text{ Mod } \Lambda, \quad (14)$$

where $\Lambda \in \mathbb{R}^T$ is the coarse lattice used in the nested lattice encoding method for the transmitted symbols $\mathbf{x}_k^l, \forall k, l$. $\{\mathbf{t}_k^l, \tilde{\mathbf{t}}_k^l\}$ are points in the nested lattice code \mathcal{L}_k^l (corresponding to the fine lattice Λ_k^l) that carry information. $\{\mathbf{d}_k^l, \tilde{\mathbf{d}}_k^l\}$ are the dither vectors [17], [18] which are independently uniformly distributed over \mathcal{V} and are available to all transmitters and receivers. Specifically, all the nested lattices $\{\Lambda_k^l\}$ are AWGN good, and the coarse lattice Λ is quantization good². In [18], such lattices and dither vectors satisfying the transmit power constraint $\frac{1}{T} \mathbb{E}[\|\mathbf{x}_k^l\|^2] = P$ are proved to exist.

Remark 1 (Complexity of the Encoding Method): The complexity of the encoding method is similar to that in the baselines. In the baseline methods, standard Gaussian random codebook is assumed to achieve the mutual information rate. On the other hand, there exists efficient coding schemes by using a scalar constellation coupled with a linear code and this could come close to the nested lattice performance (without the shaping gain) [26]. ■

Note that, dithering is a common randomization technique in lattice quantization for source coding [17]. The following two lemmas from [17], [18] capture the key properties of the dithered nested lattice codes.

Lemma 1 (Erez-Zamir [17]): For any random variable $\mathbf{t} \in \mathcal{V}$, if \mathbf{d} is statistically independent of \mathbf{t} and uniformly distributed over \mathcal{V} , $[\mathbf{t} - \mathbf{d}] \text{ Mod } \Lambda$ is uniformly distributed over \mathcal{V} , and is statistically independent of \mathbf{t} . ■

Lemma 2 (Nazer-Gastpar [18]): Let $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{T \times T})$ and \mathbf{d}_i^n be statistically independently uniformly distributed over \mathcal{V} with $P = \frac{1}{T} \mathbb{E}[\|\mathbf{d}_i^n\|^2]$, and $\mathbf{z}_k^l = \alpha \mathbf{z} + \sum_{i,n} \theta_i^n \mathbf{d}_i^n$ for some constant α and θ_i^n . The density of \mathbf{z}_k^l is upper bounded by the density of an i.i.d. zero-mean Gaussian vector $\tilde{\mathbf{z}}_k^l$ whose variance approaches N_k^l as $T \rightarrow \infty$, where $N_k^l = \alpha^2 + P \sum_{i,n} (\theta_i^n)^2$. ■

Remark 2 (Interpretation of the Lemmas): Lemma 1 assures that the input power exactly meets the power constraint [17]. Lemma 2 indicates the non-Gaussian noise \mathbf{z}_k^l is nearly Gaussian as the code length T increases. ■

B. Lattice Alignment with Imperfect CSI

In this section, we shall discuss the proposed robust lattice alignment method (using vector space strategies) and the motivations. When Gaussian signals are transmitted in interference channels, the associated interference space is random as illustrated in Fig. 2 [10]. Intuitively, the aggregate interference will fill the entire signal space and there is no room left for desired signal transmission. Furthermore, the *penalty* of interference scales with the number of interferers. As a result,

²Please refer to [16] for the definition of goodness for lattice codes. Specifically, the lattice codes should have both good statistical and good algebraic properties [18].

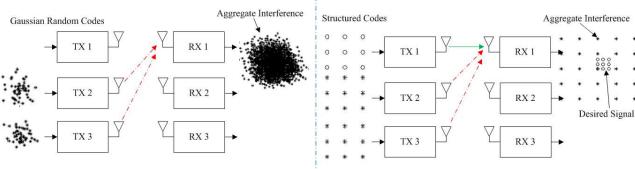


Fig. 2. Illustration of random interference (due to random codes) and structured interference (due to structured codes) in interference channels [10, Fig. 4, Fig. 5]. The resulting random interference covers the entire space, preventing the desired receiver from decoding. On the other hand, the structured interference allows the desired receiver to decode between gaps.

the key idea behind the IA methods in [3], [4] is to align all interferences to a lower dimensional subspace and utilize the remaining interference free dimensions to transmit the desired signals. However, one problem associated with the IA method in [3], [4] is the feasibility problem. In fact, it is shown in [7] that the per user DoF greater than $\frac{M+N}{K+1}$ is not achievable.

On the other hand, when structured lattices are used at the transmitters, the aggregate interference at each of the receiver may be structured (if the interfering lattices are properly aligned) as illustrated in Fig. 2 [10]. In this case, the interference space contains regular *gaps* that can be utilized to transmit desired signals. More importantly, the *penalty* of the interference does not scale with the number of interferers. Therefore, another possible direction of alignment is to align the transmit lattices on a common basis at each of the K receivers to create a structured interference space as illustrated in Fig. 2 [10]. However, it is not always possible to simultaneously align all the interfering lattices, as shown in Example 1. Furthermore, perfect lattice alignment requires perfect CSI, which is impractical.

Motivated by the advantages of having structured interference, we propose a *robust lattice alignment method* in which the precoders are designed to *align* the received lattices as much as possible. We accept the fact that lattice alignment may not be perfect (due to infeasible irrational channel coefficients and imperfect CSI) and we try to minimize the effect of the *residual lattice alignment error*. Specifically, instead of decoding and subtracting each interference individually, we wish to decode the structured aggregate interference composed of all the interferences and remove them all at once. Furthermore, due to the structural constraints of the lattice, we must choose integer coefficient for each data stream to approximate the equivalent channel coefficient. Therefore, the *aggregate interference lattice* for the data stream \mathbf{x}_k^l is written as:

$$\mathbf{I}_k^l = \sum_{i,n} a_i^n \mathbf{x}_i^n, \quad (15)$$

where $a_i^n \in \mathbb{Z} + j\mathbb{Z}$ (complex integer) is defined as the *interference quantization coefficient* to approximate the equivalent channel coefficient for \mathbf{x}_i^n , and $a_k^l = 0$ (see Fig. 3 for an illustration). Specifically, given the precoder \mathbf{v}_i^n for \mathbf{x}_i^n , the decorrelator \mathbf{u}_k^l at receiver k , and the interference quantization coefficients $\{a_i^n\}$, the alignment error is given by:

$$\mathbf{I}_e = P \sum_{i \neq k} \sum_{n \neq l} \left| (\mathbf{u}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n - a_i^n \right|^2. \quad (16)$$

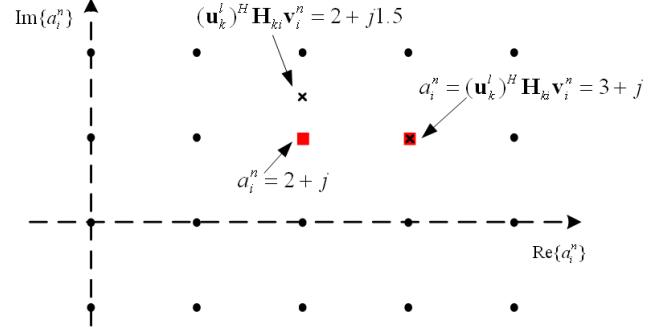


Fig. 3. Illustration of residual alignment error after stage I decoding and how it is affected by different choices of stage I decorrelator \mathbf{u}_k^l and interference quantization coefficients $\mathbf{a} = \{\{a_i^n\}_{i=1}^K\}_{n=1}^L$. Specifically, a_i^n is the interference quantization coefficient and $\mathbf{H}_{ki} \mathbf{v}_i^n$ is the equivalent channel for the stream \mathbf{x}_i^n respectively.

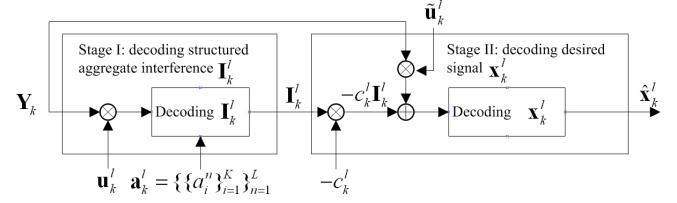


Fig. 4. The block diagram of the two-stage decoding algorithm for decoding the desired signal \mathbf{x}_k^l . Specifically, \mathbf{Y}_k is the observed signal for user k given by (19), \mathbf{I}_k^l is the structured aggregate interference given by (21), \mathbf{u}_k^l is the first stage decorrelator, \mathbf{a}_k^l are the interference quantization coefficients, c_k^l is the scaling coefficient, and $\tilde{\mathbf{u}}_k^l$ is the second stage decorrelator.

Remark 3 (The Effect of Design Parameters on \mathbf{I}_e): Note that if the channel coefficients $\{\mathbf{H}_{ki}\}_{i,k=1}^K$ are irrational, the alignment error \mathbf{I}_e may not be zero but the design parameters $\{\mathbf{u}_k^l, \mathbf{v}_i^n, a_i^n\}$ in (16) can be chosen to minimize the effect of the alignment error as illustrated in Fig. 3. We shall elaborate the precise optimization problem in Section IV-E. ■

Based on the *aggregate interference lattice* in (15), a *two-stage decoding algorithm* can be constructed at each of the receivers as illustrated in Fig. 4. The first stage is to decode the structured aggregate interference, whereas the second stage is to decode the desired signal after canceling the decoded aggregate interference. The decoding process and the associated error analysis in the presence of the residual alignment errors are discussed in detail in the following subsections.

C. Stage I - Decoding the Aggregate Interference

Note that the received signal at the k -th receiver is given by:

$$\mathbf{Y}_k = \mathbf{H}_{kk} \mathbf{v}_k^l \mathbf{x}_k^l + \sum_{n \neq l} \mathbf{H}_{kk} \mathbf{v}_k^n \mathbf{x}_k^n + \sum_{i \neq k} \sum_n \mathbf{H}_{ki} \mathbf{v}_i^n \mathbf{x}_i^n + \mathbf{Z}_k. \quad (19)$$

$$R_i^n < \mu_k^l = \begin{cases} \log \left(\frac{P}{||\mathbf{u}_k^l||^2 + P \sum_{i,n} \left| (\mathbf{u}_k^l)^H \hat{\mathbf{H}}_{ki} \mathbf{v}_i^n - a_i^n \right| + \epsilon \|\mathbf{v}_i^n\| \cdot \|\mathbf{u}_k^l\|} \right) & \text{if } a_i^n \neq 0 \\ \infty & \text{if } a_i^n = 0 \end{cases}, \quad (17)$$

$$R_k^l < \tilde{\mu}_k^l = \log \left(\frac{P}{||\tilde{\mathbf{u}}_k^l||^2 + P \sum_{i,n} \left| ((\tilde{\mathbf{u}}_k^l)^H \hat{\mathbf{H}}_{ki} \mathbf{v}_i^n - c_k^l a_i^n - 1_{\{i=k \& n=l\}}) + \epsilon \|\mathbf{v}_i^n\| \cdot \|\tilde{\mathbf{u}}_k^l\| \right|^2} \right). \quad (18)$$

At stage I, a complex linear decorrelator \mathbf{u}_k^l ($N \times 1$) is applied at receiver k , i.e.,

$$\begin{aligned} \mathbf{y}_k^l &= (\mathbf{u}_k^l)^H \mathbf{Y}_k = \underbrace{(\mathbf{u}_k^l)^H \mathbf{H}_{kk} \mathbf{v}_k^l \mathbf{x}_k^l}_{\text{desired signal}} + \underbrace{\sum_{n \neq l} (\mathbf{u}_k^l)^H \mathbf{H}_{kk} \mathbf{v}_k^n \mathbf{x}_k^n}_{\text{inter-stream interference}} \\ &\quad + \underbrace{\sum_{i \neq k; n} (\mathbf{u}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n \mathbf{x}_i^n}_{\text{inter-user interference}} + (\mathbf{u}_k^l)^H \mathbf{Z}_k. \end{aligned} \quad (20)$$

From \mathbf{y}_k^l ($1 \times T$), we wish to decode the structured aggregate interference, i.e.,

$$\mathbf{I}_k^l = \left[\sum_{i,n} \Re\{a_i^n \mathbf{x}_i^n\} \right] \text{Mod } \Lambda + j \left[\sum_{i,n} \Im\{a_i^n \mathbf{x}_i^n\} \right] \text{Mod } \Lambda, \quad (21)$$

where $a_i^n \in \mathbb{Z} + j\mathbb{Z}$, and $a_k^l = 0$. Let $\mathbf{a}_k^l = \{\{a_i^n\}_{i=1}^K\}_{n=1}^L$ denote the *interference quantization coefficients* for decoding the desired data stream \mathbf{x}_k^l . $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts, respectively. To successfully decode \mathbf{I}_k^l , there are some requirements on the transmit data rates of the streams in the \mathbf{I}_k^l term. The achievable rate region of the K users for successful stage I decoding under the imperfect CSI model in (4) is summarized in Lemma 3.

Lemma 3: (A Sufficient Condition for Successful Stage I Decoding under Imperfect CSI): Under the imperfect CSI model in (4), the structured aggregate interference \mathbf{I}_k^l can be decoded from \mathbf{y}_k^l with arbitrarily small error probability if the data rate R_i^n satisfies the condition (for all i, n) given in (17), where $\|\mathbf{v}_i^n\| \cdot \|\mathbf{u}_k^l\|$ means the product of the two terms $\{\|\mathbf{v}_i^n\|, \|\mathbf{u}_k^l\|\}$. ■

Proof: Please refer to Appendix A. Note that a similar result with perfect CSI is derived in [18] for single-stream wireless relay networks. ■

Note that when $a_i^n = 0$ for all $\{i, n\}$, i.e., $\mathbf{a}_k^l = 0$, there will be no stage I decoding for the desired data stream \mathbf{x}_k^l and the proposed method reduces to the conventional precoder-decorrelator optimization with one-stage decoding. The first term and the second term in the denominator of (17) correspond to the noise term and the residual lattice alignment errors, respectively.

D. Stage II - Decoding the Desired Signal

After successfully decoding the structured aggregate interference \mathbf{I}_k^l at stage I, we wish to decode the desired signal \mathbf{x}_k^l at stage II. This is illustrated in Fig. 4. Specifically, the scaled *structured aggregate interference* $c_k^l \mathbf{I}_k^l$ is subtracted from the received signal and the desired signal is decoded via $\tilde{\mathbf{y}}_k^l$ given

by:

$$\tilde{\mathbf{y}}_k^l = \begin{cases} (\tilde{\mathbf{u}}_k^l)^H \mathbf{Y}_k - (\tilde{\mathbf{u}}_k^l)^H \mathbf{H}_{km} \mathbf{v}_m^d \mathbf{x}_m^d & \text{If } \mathbf{a}_k^l = \delta_m^d \\ \left[\Re\{(\tilde{\mathbf{u}}_k^l)^H \mathbf{Y}_k - c_k^l \mathbf{I}_k^l\} \right] \text{Mod } \Lambda + \\ j \left[\Im\{(\tilde{\mathbf{u}}_k^l)^H \mathbf{Y}_k - c_k^l \mathbf{I}_k^l\} \right] \text{Mod } \Lambda & \text{otherwise} \end{cases}, \quad (22)$$

where $\tilde{\mathbf{u}}_k^l$ is the second stage decorrelator, $c_k^l \in \mathbb{Z} + j\mathbb{Z}$ is the *scaling coefficient*³ for the l -th data stream at receiver k , δ_m^d is the δ -function (all $a_i^n = 0$ except $a_m^d = 1$), and

$$\begin{aligned} & \left[\Re\{(\tilde{\mathbf{u}}_k^l)^H \mathbf{Y}_k - c_k^l \mathbf{I}_k^l\} \right] \text{Mod } \Lambda \\ &= \left[\Re\{(\tilde{\mathbf{u}}_k^l)^H \mathbf{Y}_k\} - \Re\{c_k^l\} \Re\{\mathbf{I}_k^l\} + \Im\{c_k^l\} \Im\{\mathbf{I}_k^l\} \right] \text{Mod } \Lambda \\ &\stackrel{(a)}{=} \left[\Re\{(\tilde{\mathbf{u}}_k^l)^H \mathbf{Y}_k\} - \left[\Re\{c_k^l\} \Re\{\mathbf{I}_k^l\} \right] \text{Mod } \Lambda \right. \\ &\quad \left. + \left[\Im\{c_k^l\} \Im\{\mathbf{I}_k^l\} \right] \text{Mod } \Lambda \right] \text{Mod } \Lambda \\ &\stackrel{(b)}{=} \left[\Re\left\{ \sum_{i,n} [(\tilde{\mathbf{u}}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n - c_k^l a_i^n] \mathbf{x}_i^n + \tilde{\mathbf{u}}_k^l \mathbf{Z}_k \right\} \right] \text{Mod } \Lambda \end{aligned} \quad (23)$$

where (a) follows from $[\mathbf{g}_1 + \mathbf{g}_2] \text{Mod } \Lambda = [([\mathbf{g}_1] \text{Mod } \Lambda) + \mathbf{g}_2] \text{Mod } \Lambda$ for all $\mathbf{g}_1, \mathbf{g}_2 \in \mathbb{R}^T$, (b) follows from $c_k^l \in \mathbb{Z} + j\mathbb{Z}$ and $[c([\mathbf{g}_1] \text{Mod } \Lambda)] \text{Mod } \Lambda = [cg_1] \text{Mod } \Lambda$ for all $c \in \mathbb{Z}$. Similarly, we have

$$\begin{aligned} & \left[\Im\{(\tilde{\mathbf{u}}_k^l)^H \mathbf{Y}_k - c_k^l \mathbf{I}_k^l\} \right] \text{Mod } \Lambda = \\ & \left[\Im\left\{ \sum_{i,n} [(\tilde{\mathbf{u}}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n - c_k^l a_i^n] \mathbf{x}_i^n + \tilde{\mathbf{u}}_k^l \mathbf{Z}_k \right\} \right] \text{Mod } \Lambda. \end{aligned} \quad (24)$$

From (22) note that, if we set $\mathbf{a}_k^l = \delta_m^d$, then $\mathbf{I}_k^l = \mathbf{x}_m^d$, and hence the stage II decoding in (22) would try to directly null off the interference \mathbf{x}_m^d . Similarly, the sufficient condition in terms of the rate region for successful stage II decoding is given in the next lemma:

Lemma 4: (A Sufficient Condition for Successful Stage II Decoding under Imperfect CSI): A sufficient condition for successful stage II decoding under the imperfect CSI model in (4) is given in (18), where $1_{\{\cdot\}}$ is an indicator function. ■

Proof: The proof follows from a similar approach as in Appendix A. Specifically, it is obtained by replacing \mathbf{y}_k^l and \mathbf{I}_k^l (obtained from (23) and (24)) with $\tilde{\mathbf{y}}_k^l$ and \mathbf{x}_k^l in the proof of Lemma 3, respectively. Note that, from (23) and (24), the equivalent channel coefficient for data stream \mathbf{x}_k^l in stage II decoding is $(\tilde{\mathbf{u}}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n - c_k^l a_i^n$. Furthermore, since we wish to decode $\mathbf{x}_k^l = 1 \cdot \mathbf{x}_k^l + \sum_{i \neq k} 0 \cdot \mathbf{x}_i^n, 1_{\{i=k \& n=l\}}$ in (18) indicates the coefficient for \mathbf{x}_k^l and the coefficients for the other data streams are all 0. ■

Note that the rate region in (18) is a function of interference quantization coefficients \mathbf{a} , scaling coefficients $\mathbf{c} =$

³ We may need to scale the decoded aggregate interference \mathbf{I}_k^l first before subtracting from the received signal so as to compensate for the amplitude change due to the second stage decorrelator $\tilde{\mathbf{u}}_k^l$.

$\{c_k^l\}_{k=1}^K\}_{l=1}^L$, precoders \mathbf{v} , and second stage decorrelators $\tilde{\mathbf{u}} = \{\tilde{\mathbf{u}}_k^l\}_{k=1}^K\}_{l=1}^L$.

Remark 4 (Complexity of the Decoding Method): The decoding complexity of the proposed design and the baselines are similar. For instance, all baseline methods have assumed ML decoding to achieve the mutual information rate. On the other hand, there exists efficient lattice decoding methods [17], [18] that could exploit the lattice symmetry in the decoding process. \blacksquare

E. Precoder, Decorrelator and Interference Quantization Coefficients Optimization

In this subsection, we shall formulate the precoders, decorrelators, scaling coefficients and interference quantization coefficients design problem formally as an optimization problem. The problem consists of the following components:

- **Optimization Variables:** Specifically, the optimization variables consist of the set of precoders \mathbf{v} , the set of stage I decorrelators \mathbf{u} , the set of stage II decorrelators $\tilde{\mathbf{u}}$, the set of interference quantization coefficients \mathbf{a} (used in stage I processing), and the set of scaling coefficients \mathbf{c} (used in stage II processing).
- **Optimization Objective:** For fairness, we consider the worst-case data rate as the optimization objective, i.e., $\max_{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{v}, \mathbf{a}, \mathbf{c}} R_{\min}$, where $R_{\min} = \min_{k,l} (\mu_k^l, \tilde{\mu}_k^l)$.
- **Optimization Constraints:** The optimization constraints consist of the transmit power constraint $\sum_{l=1}^L \|\mathbf{v}_k^l\|^2 \leq \gamma, \forall k$, the interference quantization coefficients constraint $\mathbf{a}_k^l \in (\mathbb{Z} + j\mathbb{Z})^{KL}$, and the scaling coefficients constraint $c_k^l \in \mathbb{Z} + j\mathbb{Z}$.

As a result, the optimization problem is summarized below:

$$\{\mathbf{u}^*, \tilde{\mathbf{u}}^*, \mathbf{v}^*, \mathbf{a}^*, \mathbf{c}^*\} = \begin{cases} \arg \max_{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{v}, \mathbf{a}, \mathbf{c}} \min_{l,k} (\mu_k^l, \tilde{\mu}_k^l) \\ \text{s.t. } \sum_{l=1}^L \|\mathbf{v}_k^l\|^2 \leq \gamma, \forall k \\ \mathbf{a}_k^l \in (\mathbb{Z} + j\mathbb{Z})^{KL}; c_k^l \in \mathbb{Z} + j\mathbb{Z} \end{cases} \quad (25)$$

The above optimization problem involves complex $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{v}\}$ and integer variables $\{\mathbf{a}, \mathbf{c}\}$, which is a mixed integer and continuous optimization problem. As a result, the problem is non-convex and requires exhaustive search [21].

V. LOW COMPLEXITY ITERATIVE SOLUTION

In this section, we shall propose a low complexity iterative algorithm by exploiting the special structure of the problem in (25).

A. Properties of the Optimal Interference Quantization Coefficients $\{\mathbf{a}^*\}$

Although obtaining the optimal integer solution $\{\mathbf{a}^*\}$ in (25) is a difficult problem, we have the following lemma to reduce the search space for \mathbf{a}^* .

Lemma 5 (Properties of the Optimal \mathbf{a}^):* The optimal integer solution \mathbf{a}^* in (25) should belong to the following set, i.e.,

$$(\mathbf{a}_k^l)^* \in \mathcal{A} = \{\mathbf{a}_k^l : \frac{a_k^n}{r} \notin \mathbb{Z} + j\mathbb{Z}, \forall a_k^n \neq 0, \forall r \in (\mathbb{Z} + j\mathbb{Z}) \text{ and } |r| \neq 1\} \quad \forall k, l. \quad (26)$$

Proof: Suppose $\{(\mathbf{v}_k^l)^*, (\tilde{\mathbf{u}}_k^l)^*, (\mathbf{u}_k^l)^*, (c_k^l)^*, (\mathbf{a}_k^l)^* \notin \mathcal{A}\}$ are the optimal solution. Suppose that $\frac{(a_k^n)^*}{r} \in \mathbb{Z} + j\mathbb{Z}$, and $|r| > 1$. Although \mathbf{a}_k^l will influence μ_k^l in (17) and $\tilde{\mu}_k^l$ in (18), it is easy to verify that $\mu_k^l(\frac{(a_k^n)^*}{r}, \frac{(\mathbf{a}_k^l)^*}{r}) > \mu_k^l((\mathbf{u}_k^l)^*, (\mathbf{a}_k^l)^*)$, and $\tilde{\mu}_k^l((c_k^l)^* r, \frac{(\mathbf{a}_k^l)^*}{r}) = \tilde{\mu}_k^l((c_k^l)^*, (\mathbf{a}_k^l)^*)$. In other words, a higher data rate is achievable with the designed parameters $\{(\mathbf{v}_k^l)^*, (\tilde{\mathbf{u}}_k^l)^*, (\mathbf{u}_k^l)^*, (c_k^l)^*, (\mathbf{a}_k^l)^* \notin \mathcal{A}\}$. As a result, $\{(\mathbf{v}_k^l)^*, (\tilde{\mathbf{u}}_k^l)^*, (\mathbf{u}_k^l)^*, (c_k^l)^*, (\mathbf{a}_k^l)^* \notin \mathcal{A}\}$ cannot be an optimal solution. \blacksquare

B. Optimization of $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$ under fixed $\{\mathbf{v}, \mathbf{a}\}$

In this section, we fix the precoder \mathbf{v} and interference quantization coefficients \mathbf{a} , and we optimize the remaining parameters $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$. Note that, \mathbf{u}_k^l only influences μ_k^l in (17), and $\tilde{\mathbf{u}}_k^l$ only influences $\tilde{\mu}_k^l$ in (18). For given c_k^l , we shall first determine the optimal \mathbf{u}_k^l and $\tilde{\mathbf{u}}_k^l$ by maximizing μ_k^l and $\tilde{\mu}_k^l$, respectively, which can be obtained by solving the following convex problems

$$\begin{aligned} (\mathbf{u}_k^l)^* &= \arg \min_{\mathbf{u}_k^l} \left(\|\mathbf{u}_k^l\|^2 + P \sum_{i,n} \right. \\ &\quad \left. |(\mathbf{u}_k^l)^H \hat{\mathbf{H}}_{ki} \mathbf{v}_i^n - a_i^n| + \epsilon \|\mathbf{v}_i^n\| \cdot \|\mathbf{u}_k^l\|^2 \right) \\ (\tilde{\mathbf{u}}_k^l)^* &= \arg \min_{\tilde{\mathbf{u}}_k^l} \left(\|\tilde{\mathbf{u}}_k^l\|^2 + P \sum_{i,n} \right. \\ &\quad \left. |(\tilde{\mathbf{u}}_k^l)^H \hat{\mathbf{H}}_{ki} \mathbf{v}_i^n - c_k^l a_i^n - 1_{\{i=k \& n=l\}}| + \epsilon \|\mathbf{v}_i^n\| \cdot \|\tilde{\mathbf{u}}_k^l\|^2 \right) \end{aligned} \quad (27)$$

For perfect CSI ($\epsilon = 0, \hat{\mathbf{H}} = \mathbf{H}$), closed form solutions exist for (27) and the optimal solutions are given by [27]:

$$(\mathbf{u}_k^l)^* = \left(\mathbf{W}^H \mathbf{W} + \frac{1}{P} \mathbf{I}_{N \times N} \right)^{-1} \mathbf{W}^H \boldsymbol{\alpha}_k^l, \quad (28)$$

$$(\tilde{\mathbf{u}}_k^l)^* = \left(\mathbf{W}^H \mathbf{W} + \frac{1}{P} \mathbf{I}_{N \times N} \right)^{-1} \mathbf{W}^H \boldsymbol{\beta}_k^l, \quad (29)$$

where $\mathbf{W} = [\mathbf{H}_{k1} \mathbf{v}_1^1, \dots, \mathbf{H}_{k1} \mathbf{v}_1^L, \mathbf{H}_{k2} \mathbf{v}_2^1, \dots, \mathbf{H}_{kK} \mathbf{v}_K^L]^H$ is a $KL \times N$ matrix, $\boldsymbol{\alpha}_k^l = [a_1^l, \dots, a_L^l, a_1^l, \dots, a_K^l]^H$ is a $KL \times 1$ vector, and $\boldsymbol{\beta}_k^l = [c_k^l a_1^l, \dots, c_k^l a_L^l, c_k^l a_1^l, \dots, 1, \dots, c_k^l a_K^l]^H$ is also a $KL \times 1$ vector. On the other hand, when the CSI is imperfect, there is no closed form solution. Since (27) is a standard convex problem, $(\mathbf{u}_k^l)^*$ and $(\tilde{\mathbf{u}}_k^l)^*$ can be obtained iteratively using an interior-point method or efficient gradient search [28].

Next, we shall optimize the complex integers c_k^l for a given $\tilde{\mathbf{u}}_k^l$. The solution of $(c_k^l)^*$ for given $(\tilde{\mathbf{u}}_k^l)^*$ is summarized in the lemma below.

Lemma 6 (Optimal Integer Solution of $(c_k^l)^$ given $(\tilde{\mathbf{u}}_k^l)^*$):* Given $(\tilde{\mathbf{u}}_k^l)^*$ in (27), the optimal $(c_k^l)^*$ is given by:

$$(c_k^l)^* = \arg \min_{\mathbb{R}\{c_k^l\} \in [\mathbb{R}\{\tau\}, \mathbb{R}\{\kappa\}], \Im\{c_k^l\} \in [\Im\{\tau\}, \Im\{\kappa\}]} f(c_k^l), \quad (30)$$

where

$$\tau = (c_k^l)^* - (1 + j) \text{ and } \kappa = (c_k^l)^* + (1 + j), \quad (31)$$

$$\begin{aligned} f(c_k^l) &= \\ &\sum_{i,n} \left| (\tilde{\mathbf{u}}_k^l)^H \hat{\mathbf{H}}_{ki} \mathbf{v}_i^n - c_k^l a_i^n - 1_{\{i=k \& n=l\}} \right| + \epsilon \|\mathbf{v}_i^n\| \cdot \|\tilde{\mathbf{u}}_k^l\|^2, \end{aligned} \quad (32)$$

and

$$(\tilde{c}_k^l)^* = \arg \min_{\tilde{c}_k^l \in \mathbb{C}} f(\tilde{c}_k^l). \quad (33)$$

Proof: c_k^l only influences $\tilde{\mu}_k^l$ in (18). If we relax $c_k^l \in \mathbb{Z} + j\mathbb{Z}$ to $\tilde{c}_k^l \in \mathbb{C}$, the function $f(\tilde{c}_k^l)$ is a convex function. Therefore, if $(\tilde{c}_k^l)^* = \arg \min_{\tilde{c}_k^l \in \mathbb{C}} f(\tilde{c}_k^l)$, the optimal c_k^l is one of the complex integers close to $(\tilde{c}_k^l)^*$, i.e.,

$$(c_k^l)^* = \arg \min_{\Re\{c_k^l\} \in [\Re\{\tau\}, \Re\{\kappa\}], \Im\{c_k^l\} \in [\Im\{\tau\}, \Im\{\kappa\}]} f(c_k^l), \quad (34)$$

where $\tau = (\tilde{c}_k^l)^* - (1+j)$, $\kappa = (\tilde{c}_k^l)^* + (1+j)$. \blacksquare

As a result, given $\{\mathbf{v}, \mathbf{a}\}$, we shall use the following algorithm to optimize $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$.

Subalgorithm A: Optimization Algorithm for $\{\mathbf{u}, \tilde{\mathbf{u}}\}$ and \mathbf{c} under fixed $\{\mathbf{v}, \mathbf{a}\}$

- **Step 1:** For a given realization of $\mathbf{v}, \mathbf{a}_k^l \in \mathcal{A}$, initialize $\mathbf{c}(0)$. Set the iteration steps $m = 0$.
- **Step 2:** For the given $\{\mathbf{v}, \mathbf{a}\}$, solve the convex optimization problem (27) to obtain the corresponding optimal first stage decorrelator \mathbf{u} .
- **Step 3:** For the given $\{\mathbf{v}, \mathbf{a}, \mathbf{c}(m)\}$, solve the convex optimization problem (27) to obtain the corresponding optimal second stage decorrelator $\tilde{\mathbf{u}}(m+1)$.
- **Step 4:** For the given $\{\mathbf{v}, \mathbf{a}, \tilde{\mathbf{u}}(m+1)\}$, obtain the optimal positive integers $\mathbf{c}(m+1)$ by solving the problem (30). Specifically, solve the convex problem (33) to obtain the relaxed value $(\tilde{c}_k^l)^*$, and then check the nearby integers around $(\tilde{c}_k^l)^*$ to obtain $\mathbf{c}(m+1)$ as shown in (30).
- **Step 5:** Continue until $\tilde{\mathbf{u}}(m) = \tilde{\mathbf{u}}(m+1)$, and $\mathbf{c}(m) = \mathbf{c}(m+1)$.

The convergence proof is given in Appendix B.

C. Optimization of $\{\mathbf{v}, \mathbf{a}\}$ under fixed $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$

Given $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$, optimizing $\{\mathbf{v}, \mathbf{a}\}$ is not trivial. To obtain low complexity solutions, we shall first relax the interference quantization coefficients from $\mathbf{a}_k^l \in (\mathbb{Z} + j\mathbb{Z})^{KL}$ to $\mathbf{a}_k^l \in \mathbb{C}^{KL}$. Define

$$\begin{aligned} g_k^l(\mathbf{v}, \mathbf{a}_k^l) &= \|\mathbf{u}_k^l\|^2 + P \sum_{i,n} \\ &\left| \left| (\mathbf{u}_k^l)^H \hat{\mathbf{H}}_{ki} \mathbf{v}_i^n - a_i^n \right| + \epsilon \|\mathbf{v}_i^n\| \cdot \|\mathbf{u}_k^l\| \right|^2 \\ \tilde{g}_k^l(\mathbf{v}, \mathbf{a}_k^l) &= \|\tilde{\mathbf{u}}_k^l\|^2 + P \sum_{i,n} \\ &\left| \left| (\tilde{\mathbf{u}}_k^l)^H \hat{\mathbf{H}}_{ki} \mathbf{v}_i^n - c_k^l a_i^n - 1_{\{i=k \& n=l\}} \right| + \epsilon \|\mathbf{v}_i^n\| \cdot \|\tilde{\mathbf{u}}_k^l\| \right|^2, \end{aligned}$$

where the right hand side of these two equations are the denominator of (17) and (18), respectively. $\{g_k^l, \tilde{g}_k^l\}$ are convex functions w.r.t. $\{\mathbf{v}, \mathbf{a}_k^l\}$ for all $\{l, k\}$. Given $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$, the max-min problem in (25) is equivalent to

$$\begin{aligned} \max_{\mathbf{v}, \mathbf{a}, t} \quad & t \\ \text{s.t.} \quad & t \leq \mu_k^l \text{ and } t \leq \tilde{\mu}_k^l, \quad \forall k, l \\ & \sum_{l=1}^L \|\mathbf{v}_k^l\|^2 \leq \gamma, \forall k, \end{aligned} \quad (35)$$

which is not a convex problem, because $\log(g_k^l(\mathbf{v}, \mathbf{a}_k^l))$ and $\log(\tilde{g}_k^l(\mathbf{v}, \mathbf{a}_k^l))$ are not convex functions w.r.t. $\{\mathbf{v}, \mathbf{a}_k^l\}$ for all $\{l, k\}$. However, we show that (35) is equivalent to a convex problem given in the following lemma.

Lemma 7 (Equivalent Convex Problem): The problem in (35) is equivalent to the following convex problem

$$\begin{aligned} \min_{\mathbf{v}, \mathbf{a}, t} \quad & t \\ \text{s.t.} \quad & g_k^l(\mathbf{v}, \mathbf{a}_k^l) \leq t \text{ and } \tilde{g}_k^l(\mathbf{v}, \mathbf{a}_k^l) \leq t, \forall l, k \\ & \sum_{l=1}^L \|\mathbf{v}_k^l\|^2 \leq \gamma, \forall k. \end{aligned} \quad (36)$$

Proof: We show that given $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$, the max-min problem in (25) is equivalent to

$$\arg \min_{\mathbf{v}, \mathbf{a}} \max_{l,k} (g_k^l, \tilde{g}_k^l). \quad (37)$$

First, note that given $\{\mathbf{v}, \mathbf{a}\}$, the optimal $\{l^*, k^*\}$ obtained from $\arg \max_{l,k} (g_k^l, \tilde{g}_k^l)$ is the same as $\arg \min_{l,k} (\mu_k^l, \tilde{\mu}_k^l)$. Furthermore, note that $\arg \min_{\mathbf{v}, \mathbf{a}} (g_{k^*}^{l^*}, \tilde{g}_{k^*}^{l^*}) = \arg \max_{\mathbf{v}, \mathbf{a}} (\mu_{k^*}^{l^*}, \tilde{\mu}_{k^*}^{l^*})$.

Therefore,

$$\arg \min_{\mathbf{v}, \mathbf{a}} \max_{l,k} (g_k^l, \tilde{g}_k^l) = \arg \max_{\mathbf{v}, \mathbf{a}} \min_{l,k} (\mu_k^l, \tilde{\mu}_k^l). \quad (38)$$

Similarly, we can also verify that (38) is equivalent to (36). \blacksquare

As a result, given $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$, we shall use the following algorithm to optimize $\{\mathbf{v}, \mathbf{a}\}$ by using an interior-point method [28, Chap.11].

Subalgorithm B: Optimization Algorithm for $\{\mathbf{v}, \mathbf{a}\}$ under fixed $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$

- **Step 1:** For a given realization of $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$, initialize $q > 0$, $\{\mathbf{v}_0, \mathbf{a}_0\}$, $\nu > 1$ and tolerance $\varsigma > 0$.
- **Step 2:** Solve the unconstrained convex problem $\{\mathbf{v}^*(q), \mathbf{a}^*(q)\} = \arg \min_{\mathbf{v}, \mathbf{a}} qt - \sum_{k,l} (\log(t - g_k^l(\mathbf{v}, \mathbf{a}_k^l)) + \log(t - \tilde{g}_k^l(\mathbf{v}, \mathbf{a}_k^l)))$, starting at $\{\mathbf{v}_0, \mathbf{a}_0\}$ by efficient gradient search⁴.
- **Step 3:** If $\frac{1}{q} < \varsigma$, stop and set $\{\mathbf{v}^*, \mathbf{a}^*\} = \{\mathbf{v}^*(q), \mathbf{a}^*(q)\}$, otherwise go to the next step.
- **Step 4:** Set $\{\mathbf{v}_0, \mathbf{a}_0\} = \{\mathbf{v}^*(q), \mathbf{a}^*(q)\}$,
- **Step 5:** Set $q = \nu q$, and go to the step 2.

The convergence proof is shown in Appendix B.

D. A Summary of the Overall Solution

The top-level optimization algorithm is summarized below (and illustrated in Fig. 5):

⁴ For example, using the Newton's method [28, Alg 9.5].

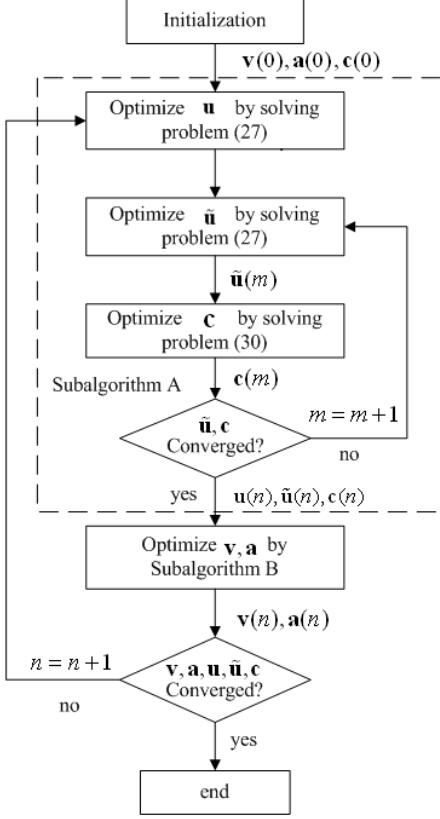


Fig. 5. Illustration of the top-level optimization algorithm. Specifically, \mathbf{v} denotes the precoders, \mathbf{a} denotes the interference quantization coefficients, \mathbf{u} denotes the first stage decorrelators, $\tilde{\mathbf{u}}$ denotes the second stage decorrelators, and \mathbf{c} denotes the scaling coefficients.

Algorithm 1: Top-Level Optimization Algorithm

- **Step 1:** Initialize $\mathbf{v}_k^l(0), \mathbf{a}_k^l(0) \in \mathcal{A}$ and $c_k^l(0)$. Set the iteration steps $m = 0$.
- **Step 2:** For the given $\{\mathbf{v}(m), \mathbf{a}(m)\}$, apply the subalgorithm A to obtain $\{\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1)\}$.
- **Step 3:** For the obtained $\{\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1)\}$, apply subalgorithm B to obtain the corresponding optimal $\{\mathbf{v}(m+1), \mathbf{a}(m+1)\}$.
- **Step 4:** Continue until $\tilde{\mathbf{u}}(m) = \tilde{\mathbf{u}}(m+1)$, $\mathbf{u}(m) = \mathbf{u}(m+1)$, $\mathbf{c}(m) = \mathbf{c}(m+1)$, $\mathbf{v}(m) = \mathbf{v}(m+1)$, and $\mathbf{a}(m) = \mathbf{a}(m+1)$.

The convergence proof of Algorithm 1 is shown in Appendix B. Finally, suppose $(\hat{\mathbf{a}}_k^l)^*$ is the solution to Algorithm 1, we could quantize $(\hat{\mathbf{a}}_k^l)^*$ to the nearest complex integers $(\hat{\mathbf{a}}_k^l)^* = \{\{(a_i^n)^*\}_{i=1}^K\}_{n=1}^L$ such that

$$(a_i^n)^* = \arg \min_{a_i^n \in \mathbb{Z} + j\mathbb{Z}} |a_i^n - (\hat{a}_i^n)^*|, \forall i, n. \quad (39)$$

Remark 5: Qualitative Comparison of the Algorithm with Conventional Approaches:

- **(a) Complexity comparisons:** The proposed method has a complexity that is close to the conventional Distributive IA method using alternating optimization [5] but as expected converges slower as seen in Table I. This is because the subproblems to be solved in each step of the

proposed algorithm (e.g. Problem (27) and Problem (36)) are convex and hence, there exists efficient algorithms (such as interior-point method [28]).

- **(b) Limitations of other existing approaches:** For the conventional IA method [3], there is always a feasibility problem for quasi-static interference channels for $K > 3$. All the other baselines (see Section VII) have poor performance when $K > 2$ even when the CSI is perfect.
- **(c) Imperfect CSI considerations:** Furthermore, none of the baselines consider imperfect CSI while the proposed method takes into account imperfect CSI in the design. Hence, they have very different performance in the presence of imperfect CSI. ■

VI. QUANTITATIVE ANALYSIS OF THE PROPOSED ROBUST LATTICE ALIGNMENT SOLUTION UNDER SYMMETRIC INTERFERENCE CHANNELS

In this section, we shall illustrate analytically the potential benefits of the proposed robust lattice alignment solution versus the brute-force two-stage ML decoding with Gaussian random inputs. To obtain a first order comparison, we consider a *symmetric interference channel* as in [11], where each user has one antenna and each transmitter tries to transmit one data stream for the desired receiver. The symmetric channel model is as follows:

$$\mathbf{y}_k = \mathbf{x}_k + h \sum_{i=1, i \neq k}^K \mathbf{x}_i + \mathbf{z}_k, \quad (40)$$

where we assume that the interference channel coefficients $h \in \mathbb{Z} + j\mathbb{Z}$. In [11], only real channel coefficients are considered.

A. Performance of the Proposed Method for Symmetric Interference Channels

Since all channels are symmetric and $M = 1$, we have $(v_k)^* = 1$ for all k , and we can consider user 1 without loss of generality. During stage I decoding, the optimal interference quantization coefficients are given by $\mathbf{a}^* = [0, 1, 1, \dots, 1]^H$ from Lemma 5. As a result, we would decode the *aggregate interference* $[\mathbf{x}_2 + \mathbf{x}_3 + \dots + \mathbf{x}_K] \bmod \Lambda$ without any quantization approximation. By solving the optimization problem in (27), we can obtain the optimal first stage decorrelator,

$$(\mathbf{u})^* = \left(\mathbf{W}^H \mathbf{W} + \frac{1}{P} \right)^{-1} \mathbf{W}^H \mathbf{a}^* = \frac{(K-1)h}{1 + (K-1)|h|^2 + 1/P}, \quad (41)$$

where $\mathbf{W} = [1, h, \dots, h]^H$, and hence we have

$$R_2^* = R_3^* = \dots = R_K^* = \log \left(\frac{1 + [(K-1)|h|^2 + 1]P}{K-1 + (K-1)P} \right). \quad (42)$$

During the second stage decoding, we have $c_k^* = h$ and $(\tilde{\mathbf{u}}_k^l)^* = 1$, where the data rate for user 1 is given by $R_1^* = \log(P)$. As a result, the minimal achievable data rate is given by:

$$R_{\min} = \min \left(\log(P), \log \left(\frac{1 + [(K-1)|h|^2 + 1]P}{K-1 + (K-1)P} \right) \right). \quad (43)$$

TABLE I
COMPARISON OF MATLAB SIMULATION TIME AND THE CONVERGED PERFORMANCE FOR A CHANNEL REALIZATION AT TRANSMIT SNR=1.5dB FOR
 $M = N = 2$ AND $L = 1$ WITH PERFECT CSI $\epsilon = 0$

Number of Users	Methods	Time (s)	Worst-Case Goodput (b/s/Hz)	Sum Goodput (b/s/Hz)
$K = 3$	Proposed Method	4.442407	1.4864	4.4593
	Distributive IA	1.579973	0.3306	2.4724
$K = 4$	Proposed Method	9.009197	0.9036	3.6144
	Distributive IA	2.995086	0.2537	2.7012

As shown from (43), the performance bottleneck of the proposed method is the stage I decoding at the medium SNR regime. Furthermore, the proposed method achieves similar performance as in [11] and hence, the proposed method is backward compatible with that in [11] under the same symmetric interference channel realizations.

B. Comparison with Brute-Force Two Stage ML Decoding with Gaussian Inputs

In this section, we shall compare the performance of the proposed method with a baseline, namely the brute-force two-stage ML decoding with random Gaussian inputs. Specifically, each transmitter transmits random Gaussian signals \mathbf{x}_k and each receiver of the K -user symmetric interference channel in (40) performs brute-force two-stage ML decoding. Without loss of generality, we consider user 1 in the analysis. During the first stage decoding, the Gaussian interferences $\{\mathbf{x}_2, \dots, \mathbf{x}_K\}$ are decoded by using brute-force ML. The decoded aggregate interference is canceled from the received signal \mathbf{y}_1 and the desired signal \mathbf{x}_1 is decoded by a second stage ML decoding. Note that unlike the proposed method, random Gaussian signals are transmitted and hence, the interference space is random. In order to have successful decoding in both stages of this system, the data rate has to satisfy the following constraints [9]:

$$\begin{aligned} R_1 &\leq \log(1+P) \\ \sum_{k=2}^K R_k &\leq \log\left(1 + \frac{(K-1)P|h|^2}{1+P}\right). \end{aligned} \quad (44)$$

Therefore, the minimal achievable data rate for all the users is given by:

$$R_{B2} = \min\left(\frac{1}{(K-1)} \log\left(1 + \frac{(K-1)P|h|^2}{1+P}\right), \log(1+P)\right), \quad (45)$$

where the bottleneck is also to decode the interference at medium SNR regime. Similarly, the conditions in which the proposed method offers a performance gain over the baseline system is given by:

$$> \frac{1}{(K-1)} \log\left(1 + \frac{(K-1)|h|^2 P}{1+P}\right). \quad (46)$$

The condition reduces to $|h|^2 \geq (1 + \frac{1}{P})$ for large K and $|h|^2 \geq \frac{(K-1)^{K-2}-1}{K-1}$ for medium SNR. In other words, only mild conditions are needed for a performance gain over the baseline system.

VII. PERFORMANCE OF PROPOSED METHOD UNDER K -USER INTERFERENCE CHANNELS

A. Comparison with Conventional Interference Alignment Solution

As pointed out in Section III-A, conventional IA can achieve $\frac{3M}{2}$ DoF in the 3-user quasi-static MIMO interference channels with $M > 1$ antennas for all transmitters and receivers [3]. However, it requires full rank channel matrices whose coefficients are randomly drawn from a continuous distribution⁵, as well as the feasibility condition. On the other hand, our proposed method works for any channel realizations and does not have feasibility problem. In this section, we shall illustrate that the proposed method could achieve the same DoF as the conventional IA [3] if it is feasible. Without loss of generality, consider the case when M is even and denote the solution of the alignment method in [3] by

$$\begin{aligned} \mathbf{V}_k(\text{IA}) &= [\mathbf{v}_k^1(\text{IA}), \dots, \mathbf{v}_k^{M/2}(\text{IA})], \\ \mathbf{U}_k(\text{IA}) &= [\mathbf{u}_k^1(\text{IA}), \dots, \mathbf{u}_k^{M/2}(\text{IA})], \end{aligned} \quad (47)$$

where $\mathbf{v}_k^l(\text{IA})$ and $\mathbf{u}_k^l(\text{IA})$ are the precoder and decorrelator for the l -th data stream of user k respectively, such that

$$\begin{aligned} \mathbf{u}_k^l(\text{IA}) \mathbf{H}_{kk} \mathbf{v}_k^l(\text{IA}) &\neq 0; \\ \mathbf{u}_k^l(\text{IA}) \mathbf{H}_{ki} \mathbf{v}_i^n(\text{IA}) &= 0, \forall i \neq k \text{ or } \forall n \neq l. \end{aligned} \quad (48)$$

The same DoF of $\frac{3M}{2}$ can be achieved in our proposed method by setting $L = \frac{M}{2}$ and choosing the design parameters as

$$\begin{aligned} \mathbf{a}_k^l &= \mathbf{0}; \mathbf{u}_k^l = \mathbf{0}; c_k^l = 1; \mathbf{v}_k^l = \mathbf{v}_k^l(\text{IA}); \\ \tilde{\mathbf{u}}_k^l &= \frac{P\zeta}{\|\mathbf{u}_k^l(\text{IA})\|^2 + P|\zeta|^2} \mathbf{u}_k^l(\text{IA}), \end{aligned} \quad (49)$$

where $\zeta = (\mathbf{u}_k^l(\text{IA}))^H \mathbf{H}_{kk} \mathbf{v}_k^l(\text{IA}) \neq 0$. Since $\mathbf{a}_k^l = \mathbf{0}$, there is no stage I decoding and hence, from (17), the data rate constraint on the data stream is ∞ . This means that there is no data rate constraint associated with stage I decoding. As a result, the desired data stream \mathbf{x}_k^l is decoded directly at stage II. From (18) and (49), the achievable data rate R_k^l for data stream \mathbf{x}_k^l is given by:

$$R_k^l < \log\left(1 + \frac{P|\zeta|^2}{\|\mathbf{u}_k^l(\text{IA})\|^2}\right). \quad (50)$$

Therefore, the DoF $3L = \frac{3M}{2}$ can also be achieved in our proposed method. Hence, our proposed method achieves the same DoF performance as the conventional interference alignment method in [3], [4], as long as the problem is feasible. On the other hand, when the conventional alignment method is infeasible, our proposed solution still offers significant performance gains compared with various baseline systems.

⁵Therefore, the special cases are not considered in conventional IA. e.g., key hole effect, line-of-sight and closely spaced antennas.

B. Numerical Simulations

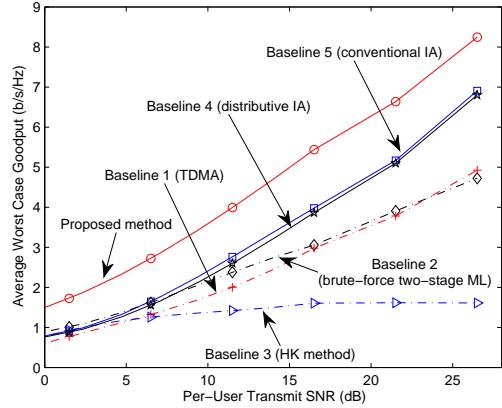
In this section, we shall compare the system goodput (b/s/Hz successfully received) of the proposed robust lattice alignment method (designed to optimize the worst-case data rate) with five baselines. Baseline 1 refers to the TDMA method. Baseline 2 refers to the brute-force two-stage ML decoding with Gaussian inputs described in Section VI-B. Baseline 3 refers to the generalized HK method for $K > 2$ users, where each user's message is split into one private part and several common parts as described in [12]. Baseline 4 refers to distributive IA based on the alternating optimization method in [5], which tries to minimize the sum leakage interference. Baseline 5 refers to the conventional IA method in [3]. In the simulation, we assume that all the channel coefficients are i.i.d. zero mean unit variance circularly symmetric complex Gaussian.

Fig. 6 and Fig. 7 illustrate the average worst-case goodput and the sum goodput versus transmit SNR (dB), respectively, at CSI errors $\epsilon = \{0, 0.1\}$ for $K = 3$, $M = N = 2$ and $L = 1$. It can be observed that our proposed method outperforms all the baselines under both perfect and imperfect CSI. Furthermore, the conventional IA and distributive IA methods achieve good performance with perfect CSI but they are very sensitive to CSI errors, especially in the high SNR region. On the other hand, our proposed method is robust to CSI errors and outperforms all the baselines as illustrated in Fig. 6. Note that our optimization objective is chosen to be the worst-case data rate for fairness consideration, which may cause some penalty⁶ on the system sum rate (as illustrated in Fig. 7(a)).

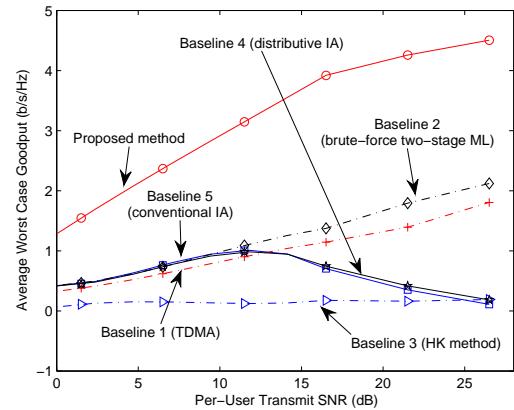
Fig. 8 illustrates the average worst-case goodput versus transmit SNR (dB) with CSI errors $\epsilon = \{0, 0.1\}$ for $K = 4$, $M = N = 2$ and $L = 1$. Note that in this configuration, the conventional IA method [3] is not feasible but the proposed method could achieve significant performance gain over all the other baselines.

Fig. 9 illustrates the average sum and worst-case goodput versus CSI error ϵ at transmit SNR=11.5dB for $K = 3$, $M = N = 2$ and $L = 1$. It can be observed that all the baselines are very sensitive to imperfect CSI. Even a small CSI error will result in a significant degradation in the system performance. On the other hand, our method is quite robust to imperfect CSI. Fig. 10 illustrates the cumulative distribution function (CDF) of the worst-case instantaneous mutual information per user at the transmit SNR=11.5dB for $K = 4$, $M = N = 4$ and $L = 2$ with CSI error $\epsilon = 0.1$. Observe that our proposed method still offers significant performance gain in the outage sense.

Fig. 11 illustrates the average worst-case goodput per user versus the number of users K at transmit SNR=1.5dB for $M = N = 4$ and $L = 2$ with CSI error $\epsilon = 0.1$. Observe that the performance degrades as K increases, and it is more difficult to decode the desired signal. Nevertheless, the proposed method outperforms all the baselines at all K . Finally, we compare the performance of the proposed method at low SNR



(a) Perfect CSI (CSI error $\epsilon = 0$)



(b) Imperfect CSI (CSI error $\epsilon = 0.1$)

Fig. 6. Comparison of average worst-case goodput (b/s/Hz successfully received) versus transmit SNR (dB) with CSI errors $\epsilon = \{0, 0.1\}$. The setup is given by $K = 3$ (number of users), $M = N = 2$ (transmit and receive antennas) and $L = 1$ (number of data stream).

with a naive method that treats interference as noise⁷ in Fig. 12. Observe that the proposed method outperforms the naive method at low SNR and the gain is contributed by precoding.

VIII. CONCLUSION

In this paper, we proposed a robust lattice alignment design for K -user quasi-static MIMO interference channels. We exploit the *structured interference* and propose a robust lattice alignment method for irrational interference channels with imperfect CSI at all SNR regimes as well as a two-stage decoding algorithm to decode the desired signal from the structured interference space. For fairness, we maximize the worst-case data rate by formulating the design of the precoders, decorrelators, scaling coefficients and the interference quantization coefficients as a mixed integer and continuous optimization problem. By using an alternating optimization technique and by incorporating imperfect CSI in the design, we derive robust solutions with low complexity. Numerical

⁶Had we changed the optimization objective to sum rate, the proposed method would also outperform the others in Fig. 7(a).

⁷It is shown in [2] that treating interference as noise is a reasonable strategy at low SNR. Furthermore, the naive method does not require CSI at the transmitter and is a robust method.

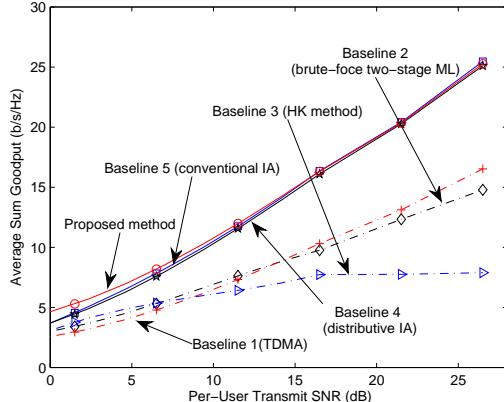
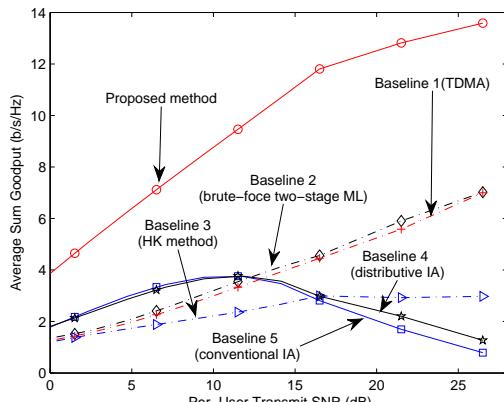
(a) Perfect CSI (CSI error $\epsilon = 0$)(b) Imperfect CSI (CSI error $\epsilon = 0.1$)

Fig. 7. Comparison of the average sum goodput (b/s/Hz successfully received) versus transmit SNR (dB) with CSI errors $\epsilon = \{0, 0.1\}$. The setup is the same as that of Fig. 6.

results verify the advantages of the proposed robust lattice alignment method compared with various baseline systems.

APPENDIX A PROOF OF LEMMA 3

From (21), the real part of the structured aggregate interference is given by:

$$\begin{aligned} \Re\{\mathbf{I}_k^l\} &= \left[\sum_{i,n} \Re\{a_i^n \mathbf{x}_i^n\} \right] \text{mod } \Lambda \\ &= \left[\left(\sum_{i,n} \Re\{a_i^n\} \mathbf{t}_i^n - \Im\{a_i^n\} \tilde{\mathbf{t}}_i^n \right) \text{Mod } \Lambda - \right. \\ &\quad \left. \sum_{i,n} \Re\{a_i^n\} \mathbf{d}_i^n - \Im\{a_i^n\} \tilde{\mathbf{d}}_i^n \right] \text{Mod } \Lambda. \end{aligned} \quad (51)$$

Thus, decoding $\Re\{\mathbf{I}_k^l\}$ is equivalent to decoding $\mathbf{T}_k^l = \left[\sum_{i,n} \left(\Re\{a_i^n\} \mathbf{t}_i^n - \Im\{a_i^n\} \tilde{\mathbf{t}}_i^n \right) \right] \text{Mod } \Lambda$. Note that the rate of \mathbf{I}_k^l is determined by the maximal rate of $\mathcal{L}_i^n, \forall a_i^n \neq 0$. Without loss of generality, the maximal data rate is denoted as R_{\max} and the corresponding nested lattice is given by Λ_{\max} . The

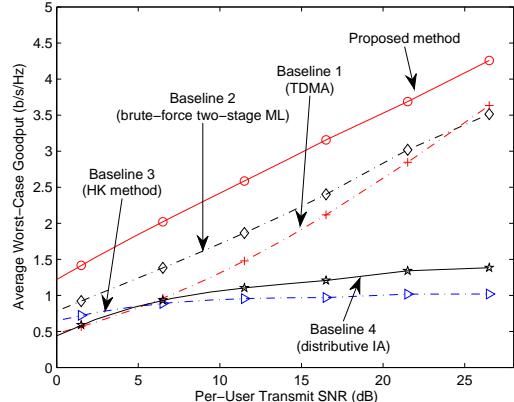
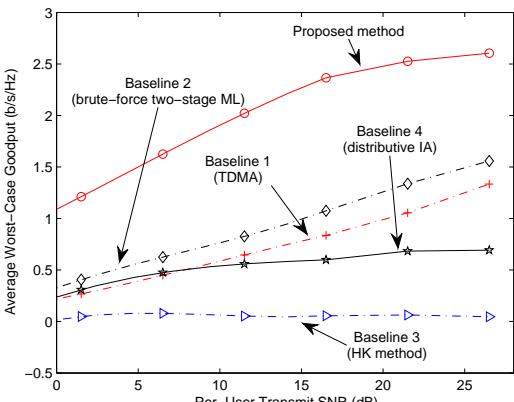
(a) Perfect CSI (CSI error $\epsilon = 0$)(b) Imperfect CSI (CSI error $\epsilon = 0.1$)

Fig. 8. Comparison of average worst-case goodput (b/s/Hz successfully received) versus transmit SNR (dB) with CSI errors $\epsilon = \{0, 0.1\}$. The setup is given by $K = 4$ (number of users), $M = N = 2$ (transmit and receive antennas) and $L = 1$ (number of data stream). Note that in this configuration, the conventional IA method is not feasible [7].

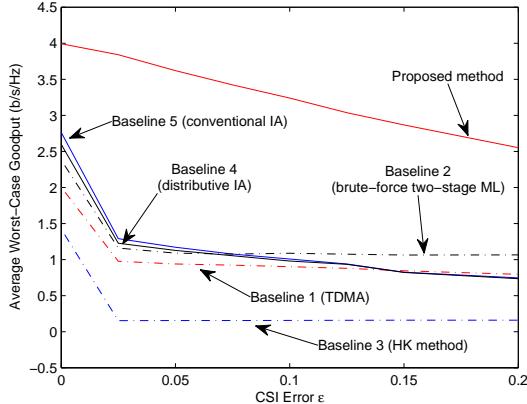
estimate of the \mathbf{T}_k^l at receiver k is given by:

$$\begin{aligned} \hat{\mathbf{T}}_k^l &= \left[Q_{\Lambda_{\max}} \left(\Re\{\mathbf{y}_k^l\} + \sum_{i,n} \Re\{a_i^n\} \mathbf{d}_i^n - \Im\{a_i^n\} \tilde{\mathbf{d}}_i^n \right) \right] \text{Mod } \Lambda \\ &= \left[Q_{\Lambda_{\max}} \left(\mathbf{T}_k^l + \mathbf{z}_k^l \right) \right] \text{Mod } \Lambda, \end{aligned} \quad (52)$$

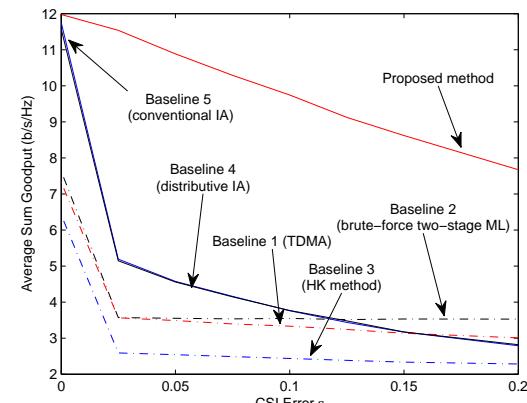
where $\mathbf{z}_k^l = \Re\{\mathbf{u}_k^l \mathbf{Z}_k\} + \sum_{i,n} \Re\{(\mathbf{u}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n - a_i^n\} \Re\{\mathbf{x}_i^n\} - \Im\{(\mathbf{u}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n - a_i^n\} \Im\{\mathbf{x}_i^n\}$. From Lemma 1, $\Re\{\mathbf{x}_i^n\}$ and $\Im\{\mathbf{x}_i^n\}$ are independently uniformly distributed over \mathcal{V} with $\frac{1}{T} \mathbb{E} \|\Re\{\mathbf{x}_i^n\}\|^2 = \frac{1}{T} \mathbb{E} \|\Im\{\mathbf{x}_i^n\}\|^2 = P/2$, and $\mathbb{E}[\mathbf{x}_i^n] = 0$. From Lemma 2, the density is upper bounded by the density of an i.i.d. zero mean Gaussian vector $\tilde{\mathbf{z}}_k^l$ whose variance $(\sigma_k^l)^2$ approaches

$$N_k^l = \frac{\|\mathbf{u}_k^l\|^2}{2} + \frac{P}{2} \sum_{i,n} \left| (\mathbf{u}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n - a_i^n \right|^2. \quad (53)$$

We set the volume of \mathcal{V}_{\max} as $\text{Vol}(\mathcal{V}_{\max}) > (2\pi\sigma_k^l)^{T/2}$, since Λ_k^l is AWGN good, the probability of $\Pr\{\hat{\mathbf{T}}_k^l \neq \mathbf{T}_k^l\}$ goes to zero exponentially in T [16]–[18]. Let $G(\Lambda)$ denote the normalized second moment of lattice Λ , then $\text{Vol}(\mathcal{V}) =$



(a) Average worst-case goodput



(b) Average sum goodput

Fig. 9. Comparison of average sum and worst-case goodput (b/s/Hz successfully received) versus CSI error ϵ at transmit SNR=11.5dB. The setup is given by $K = 3$ (number of users), $M = N = 2$ (transmit and receive antennas) and $L = 1$ (number of data stream).

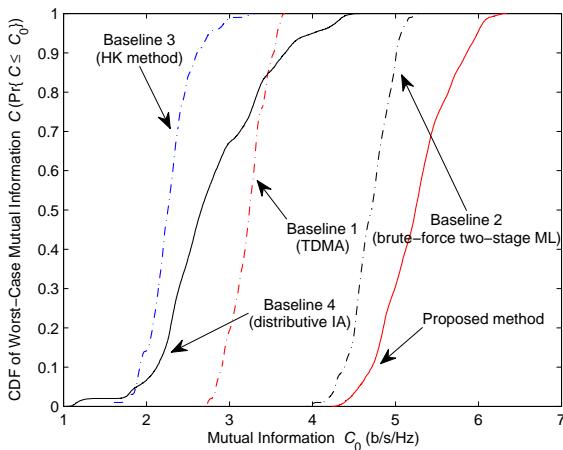


Fig. 10. The cumulative distribution function (CDF) of the instantaneous worst-case mutual information per user at the transmit SNR=11.5dB for $K = 4$, $M = N = 4$ and $L = 2$ with CSI error $\epsilon = 0.1$.

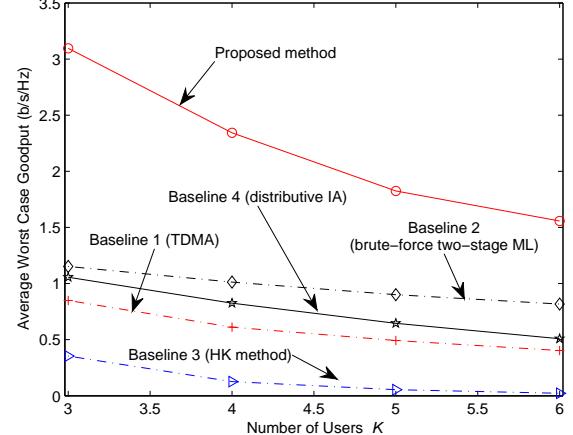


Fig. 11. The average worst-case goodput per user versus the number of users K at transmit SNR=1.5dB for $M = N = 4$ and $L = 2$ with CSI error $\epsilon = 0.1$.

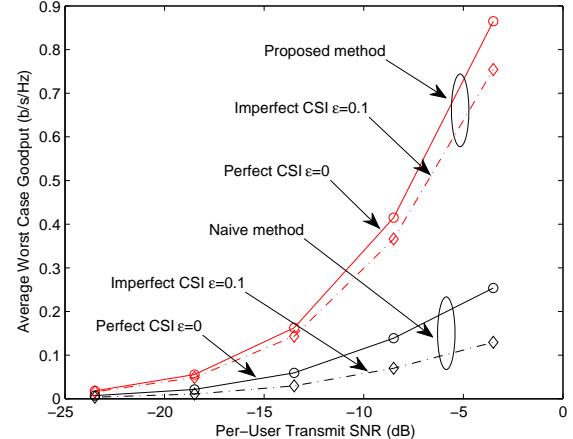


Fig. 12. Illustration of worst-case goodput (b/s/Hz successfully received) at low SNR with CSI errors $\epsilon = \{0, 0.1\}$ for $K = 4$, $M = N = 2$ and $L = 1$. The naive method refers to treating interference as noise without requiring any CSI at the transmitter. At low SNR, treating interference as noise is a reasonable strategy [2] and is also a robust method as no CSIT is needed.

$\left(\frac{P/2}{G(\Lambda)}\right)^{T/2}$ [16]. From (13), the rate of the nested lattice code \mathcal{L}_{\max} is given by:

$$R_{\max} = \frac{1}{T} \log \left(\frac{\text{Vol}(\mathcal{V})}{\text{Vol}(\mathcal{V}_{\max})} \right) < \frac{1}{2} \log \left(\frac{P/2}{G(\Lambda)2\pi e(\sigma_k^l)^2} \right). \quad (54)$$

Furthermore, for arbitrarily small $\xi > 0$, and large enough T , $G(\Lambda)2\pi e < (1 + \xi)$ [16]. By Lemma 2, if T is large enough, $(\sigma_k^l)^2 < (1 + \xi)N_k^l$. As a result, the following rate is achievable for R_{\max} :

$$\frac{1}{2} \log \left(\frac{P}{\| \mathbf{u}_k^l \|^2 + P \sum_{i,n} \| (\mathbf{u}_k^l)^H \mathbf{H}_{ki} \mathbf{v}_i^n - a_i^n \|^2} \right) - 2 \log(1 + \xi). \quad (55)$$

By choosing ξ as sufficiently small, we can neglect the second term, and the first term can be approached as desired. Similarly, $\Im\{\mathbf{I}_k^l\}$ can also be successfully decoded from $\Im\{\mathbf{y}_k^l\}$ with the same achievable data rate in (55) for R_{\max} . Note that, if $\mathbf{a}_k^l = \delta_m^d$ (all $a_i^n = 0$ except that $a_m^d = 1$),

$\mathbf{T}_k^l = [\mathbf{t}_m^d] \text{ Mod } \Lambda = \mathbf{t}_m^d$, and $[\tilde{\mathbf{t}}_m^d] \text{ Mod } \Lambda = \tilde{\mathbf{t}}_m^d$. Therefore, \mathbf{x}_m^d can be decoded in stage I and completely nulled out at stage II.

Since the data rate R_i^n is twice the lattice codes \mathcal{L}_i^n (due to the real and the imaginary parts), the minimal achievable data rate of R_i^n due to the effect of the CSI error is given by

$$R_{\min} = \min_{\Delta_{ki} \in \mathcal{E}} \log \left(\frac{P}{\|\mathbf{u}_k^l\|^2 + P \sum_{i,n} |(\mathbf{u}_k^l)^H (\tilde{\mathbf{H}}_{ki} - \Delta_{ki}) \mathbf{v}_i^n - a_i^n|^2} \right). \quad (56)$$

For given $\{\mathbf{a}_k^l, \mathbf{u}_k^l, \mathbf{v}_k^l\}$, a lower bound of R_{\min} is to choose Δ_{ki} maximizing each term of the sum in the denominator, which is given by:

$$\begin{aligned} & \max_{\Delta_{ki} \in \mathcal{E}} |(\mathbf{u}_k^l)^H (\tilde{\mathbf{H}}_{ki} - \Delta_{ki}) \mathbf{v}_i^n - a_i^n|^2 \\ \Rightarrow & \max_{\Delta_{ki} \in \mathcal{E}} |(\mathbf{u}_k^l)^H \Delta_{ki} \mathbf{v}_i^n|^2, \end{aligned} \quad (57)$$

where

$$\begin{aligned} & |(\mathbf{u}_k^l)^H \Delta_{ki} \mathbf{v}_i^n|^2 = \text{Tr}\{(\mathbf{u}_k^l)^H \Delta_{ki} \mathbf{v}_i^n (\mathbf{v}_i^n)^H \Delta_{ki}^H \mathbf{u}_k^l\} \\ = & \text{Tr}\{\mathbf{u}_k^l (\mathbf{u}_k^l)^H \Delta_{ki} \mathbf{v}_i^n (\mathbf{v}_i^n)^H \Delta_{ki}^H\} \\ \leq & \text{Tr}\{\mathbf{u}_k^l (\mathbf{u}_k^l)^H\} \text{Tr}\{\Delta_{ki} \mathbf{v}_i^n (\mathbf{v}_i^n)^H \Delta_{ki}^H\} \\ = & \text{Tr}\{\mathbf{u}_k^l (\mathbf{u}_k^l)^H\} \text{Tr}\{\mathbf{v}_i^n (\mathbf{v}_i^n)^H \Delta_{ki}^H \Delta_{ki}\} \\ \leq & \text{Tr}\{\mathbf{u}_k^l (\mathbf{u}_k^l)^H\} \text{Tr}\{\mathbf{v}_i^n (\mathbf{v}_i^n)^H\} \text{Tr}\{\Delta_{ki}^H \Delta_{ki}\} \\ = & \|\mathbf{u}_k^l\|^2 \|\mathbf{v}_i^n\|^2. \end{aligned} \quad (58)$$

Therefore, R_{\min} given in (56) is lower bounded by:

$$R_{\min} \geq \mu_k^l = \log \left(\frac{P}{\|\mathbf{u}_k^l\|^2 + P \sum_{i,n} |(\mathbf{u}_k^l)^H \tilde{\mathbf{H}}_{ki} \mathbf{v}_i^n - a_i^n| + \epsilon \|\mathbf{v}_i^n\| \cdot \|\mathbf{u}_k^l\|} \right), \quad (59)$$

which shows that $R_i^n < \mu_k^l$ is the sufficient condition for the achievable data rate R_i^n under imperfect CSI.

The proof of the Lemma 3 is an extension of the results in [18]. However, there are the following differences. Firstly, the results in [18] are for perfect CSI while our results are an extension of [18] for imperfect CSI. Secondly, the results in [18] consider single stream relay systems while our results are extended to multi-stream MIMO interference channels. Finally, the results in [18] are analysis based whereas our solution is *optimization based* where we shall optimize the precoders $\mathbf{v} = \{\{\mathbf{v}_k^l\}_{k=1}^K\}_{l=1}^L$, first stage decorrelators $\mathbf{u} = \{\{\mathbf{u}_k^l\}_{k=1}^K\}_{l=1}^L$, the scaling coefficients $\mathbf{c} = \{\{c_k^l\}_{k=1}^K\}_{l=1}^L$ and the interference quantization coefficients $\mathbf{a} = \{\{a_k^l\}_{k=1}^K\}_{l=1}^L$. The detailed formulation is illustrated in Section IV-E.

APPENDIX B

CONVERGENCE PROOF OF SUBALGORITHMS A AND B, AND TOP ALGORITHM 1

We first provide the convergence proof of subalgorithm A. Note that Given $\{\mathbf{v}, \mathbf{a}\}$, after each iteration in subalgorithm A (the alternating optimization between $\tilde{\mathbf{u}}_k^l$ and c_k^l), the data rate $\tilde{\mu}_k^l$ in (18) is increasing. Specifically, we have $\tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m+1), c_k^l(m)) > \tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m), c_k^l(m))$ in step 3 of subalgorithm A, and we have $\tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m+1), c_k^l(m+1)) > \tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m+1), c_k^l(m))$ in step 4 of subalgorithm B. Therefore,

$$\begin{aligned} & \tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m+1), c_k^l(m+1)) \\ & > \tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m+1), c_k^l(m)) > \tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m), c_k^l(m)). \end{aligned} \quad (60)$$

As a result, subalgorithm A converges. In addition, we have formulated subalgorithm B to a standard interior-point algorithm, i.e., the barrier method [28]. Therefore the results of subalgorithm B $\{\mathbf{v}^*, \mathbf{a}^*\}$ converge to a ς -optimal solution for fixed $\{\mathbf{u}, \tilde{\mathbf{u}}, \mathbf{c}\}$, and are subject to the tolerance of ς [28].

Note that $R_{\min} = \min_{k,l} (\mu_k^l, \tilde{\mu}_k^l)$, and by using the above results we have $\mu_k^l(\mathbf{u}_k^l(m+1), \mathbf{v}(m), \mathbf{a}(m)) > \mu_k^l(\mathbf{u}_k^l(m), \mathbf{v}(m), \mathbf{a}(m))$, and $\tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m+1), c_k^l(m+1), \mathbf{v}(m), \mathbf{a}(m)) > \tilde{\mu}_k^l(\tilde{\mathbf{u}}_k^l(m), c_k^l(m), \mathbf{v}(m), \mathbf{a}(m))$ in step 2 of top algorithm 1, in other words, $R_{\min}(\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1), \mathbf{v}(m), \mathbf{a}(m)) > R_{\min}(\mathbf{u}(m), \tilde{\mathbf{u}}(m), \mathbf{c}(m), \mathbf{v}(m), \mathbf{a}(m))$. Furthermore, we have $t(\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1), \mathbf{v}(m+1), \mathbf{a}(m+1)) < t(\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1), \mathbf{v}(m), \mathbf{a}(m))$ in step 3 of top algorithm 1, where t is given in (36). In other words, $R_{\min}(\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1), \mathbf{v}(m+1), \mathbf{a}(m+1)) > R_{\min}(\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1), \mathbf{v}(m), \mathbf{a}(m))$. Therefore we can conclude that

$$\begin{aligned} & R_{\min}(\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1), \mathbf{v}(m+1), \mathbf{a}(m+1)) \\ & > R_{\min}(\mathbf{u}(m+1), \tilde{\mathbf{u}}(m+1), \mathbf{c}(m+1), \mathbf{v}(m), \mathbf{a}(m)) \\ & > R_{\min}(\mathbf{u}(m), \tilde{\mathbf{u}}(m), \mathbf{c}(m), \mathbf{v}(m), \mathbf{a}(m)). \end{aligned} \quad (61)$$

Since objective R_{\min} increases after each iteration as indicated in (61), Algorithm 1 shall converge.

ACKNOWLEDGMENT

The authors would like to thank F. McKay at the Hong Kong University of Science and Technology, and the anonymous reviewers for their helpful comments and suggestions that significantly improved the quality and the presentation of the paper.

REFERENCES

- [1] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, pp. 49–60, Jan. 1981.
- [2] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5534–5562, Dec. 2008.
- [3] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3425–3441, Aug. 2008.
- [4] T. Gou and S. Jafar, "Degrees of freedom of the K user MxN MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 56, pp. 6040–6057, Dec. 2010.
- [5] K. S. Gomadam, V. R. Cadambe, and S. A. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," 2008, available: <http://arxiv.org/abs/0803.3816>.
- [6] S. W. Peters and R. W. Heath, "Interference alignment via alternating minimization," in *Proc. ICASSP09*, Apr. 2009.
- [7] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, pp. 4771–4782, Sep. 2010.
- [8] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3936–3964, Sep. 2006.
- [9] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Second Edition)*. John Wiley & Sons, Inc., 2006.
- [10] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4566–4592, Sep. 2010.
- [11] S. Sridharan, A. Jafarian, S. Vishwanath, and S. A. Jafar, "Capacity of symmetric K-user Gaussian very strong interference channels," Proceedings of IEEE GLOBECOM 2008.

- [12] S. Sridharan, A. Jafarian, S. Vishwanath, S. A. Jafar, and S. Shamai, "A layered lattice coding scheme for a class of three user Gaussian interference channels," Proceedings of 46th Annual Allerton Conference on Communication, Control and Computing, Sep 2008.
- [13] A. S. Motahari, S. O. Gharan, M. Maddaha-Ali, and A. K. Khandani, "Real interference alignment: exploiting the potential of single antenna systems," 2009. available: <http://arxiv.org/abs/0908.2282>.
- [14] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "Interference alignment for the K user MIMO interference channel," in *ISIT 2010, Austin, Texas, U.S.A.*, June 2010.
- [15] M. Maddaha-Ali, "On the degrees of freedom of the compound MIMO broadcast channels with finite states," 2009. available: <http://arxiv.org/abs/0909.5006>.
- [16] R. Zamir, "Lattices are everywhere," in *in 4th Annual Workshop on Information Theory and its Applications, UCSD, (La Jolla, CA)*, Feb. 2009.
- [17] U. Erez and R. Zamir, "Achieving $\frac{1}{2} \log(1 + \text{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2293–2314, Oct. 2004.
- [18] B. Nazer and M. Gastpar, "Compute-and-forward: harnessing interference through structured codes," 2009. available: <http://arxiv.org/abs/0908.2119>.
- [19] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath, "Ergodic interference alignment," in *Proc. ISIT*, June-July 2009.
- [20] R. Tresch and M. Guillaud, "Cellular interference alignment with imperfect channel knowledge," in *Proc. ICC*, June 2009.
- [21] D. Li, *Nonlinear Integer Programming*. New York: Springer, 2006.
- [22] J. h. Wang and D. P. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge," *IEEE Trans. Signal Process.*, vol. 57, pp. 3086–3100, Aug. 2009.
- [23] A. Abdel-Samad, T. N. Davidson, and A. B. Gershman, "Robust transmit eigen beamforming based on imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 54, pp. 1596–1609, May 2006.
- [24] A. Pascual-Iserte, D. P. Palomar, A. I. Prez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with partial channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, pp. 346–360, Jan. 2006.
- [25] H. A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1767–1773, Nov. 1997.
- [26] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3417–3432, Oct. 2005.
- [27] J. N. Zhan, B. Nazer, M. Gastpar, and U. Erez, "MIMO compute-and-forward," in *Proc. ISIT*, June-July 2009.
- [28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.